R. K. Oswald

Design of a Disk File Head-Positioning Servo

Abstract: The engineering design of a head-positioning system for an interchangeable-medium disk file is considered. Emphasis is placed upon three specific functions within the positioning system: (1) encoding and demodulation of information from the dedicated servo surface, (2) compensation and dynamics of the track-following control system, and (3) implementation of control electronics for a quasi-time-optimal, track-accessing control system. The examples used are taken from the IBM 3340 Disk Storage Facility.

Introduction

The two major functions of the head-positioning servomechanism in a disk file are track accessing and track following. The track-accessing function provides minimum time movement of the recording head from its existing track, which is a circular recording band at a specified disk radius, to a different track specified by the file controller. The track-following function maintains the position of the recording heads exactly over the center of a given track with minimum displacement error in the presence of any possible disturbances. However, with the overall system design used in a modern disk file, the servo must also provide other functions. It must furnish a clock signal, synchronized with disk rotation, to the recording channel so that all tracks will contain the same number of bits in the presence of spindle speed variations. The servo must also provide an index mark and rotational position information for the data channel, as well as a means of recalibrating the access mechanism to data track 0 on startup or after errors. These secondary requirements, as well as the major functions, have a considerable influence upon the choice of encoding used on the servo disk, as is described in the following section.

A major challenge in the design of the track-following servo is to obtain maximum bandwidth so as to minimize both settling time and displacement error but still guarantee absolute stability and freedom from oscillations. The design approach used to achieve these objectives in the IBM 3340 track-following servo is described in the third section.

The system design used for the track-accessing servo of the 3340 is not significantly different from that used in other disk files or even in other devices where the objective is minimum time motion to a specified location subject to constraints on available power and control system complexity [1]. The fourth section, however, describes some innovations in the implementation of several functions of the track-accessing servo that provide improved performance at reduced cost.

The objective of this paper is not to exhaustively cover either the theoretical bases or the analytical design procedures applicable to this work. If they are unfamiliar to the reader, he may consult his favorite texts on dynamics and feedback control, and references such as [1] and [2]. Rather, the intention here is to show some of the engineering criteria used in the choice of an encoding method for the servo disk and a compensator for the track-following servo, as well as to present several servo system design innovations that may be of value in other applications.

Servo surface encoding and demodulation

To accomplish the assigned function, a head-positioning system requires an error signal that indicates the radial position of the heads relative to the desired recording track location on the data disks. In present-day high-performance disk files such as the IBM 3340, this is the primary function of the prerecorded servo disk and read-only servo head (shown in Fig. 1) and the associated servo electronic circuits. They are examined here in detail.

To meet the requirements for a stable and accurate feedback control system, the error signal transducer must provide a signal that is not only zero when the head is properly positioned "on track", but is also linearly proportional to the size and polarized as to the direction of misposition distance when "off track". The transducer transfer function, expressed as output volts per

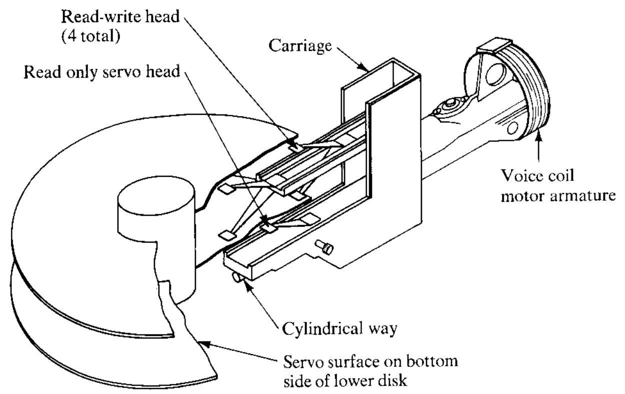

Figure 1 Mechanical components of the head-positioning system used in the IBM 3340. Two read-write heads read the top of the lower disk and two other heads on the same arm read the bottom of the upper disk. The read-only servo head utilizes the lower side of the bottom disk. The carriage rides on the cylindrical way.

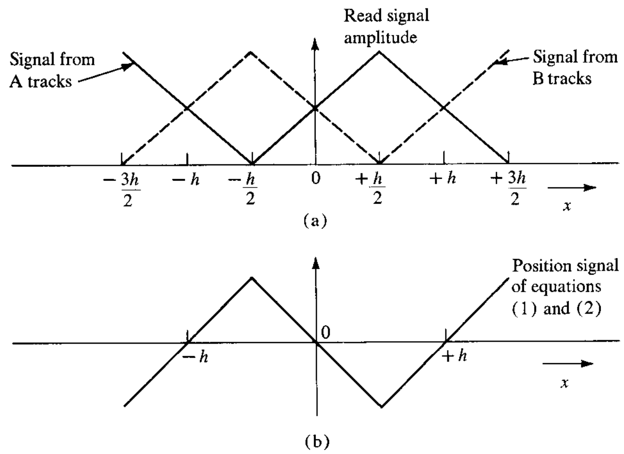

Figure 3 Form of servo head signals. (a) Amplitude of read signals from "A" and "B" tracks. (b) Position signal, described by Eqs. (1) and (2).

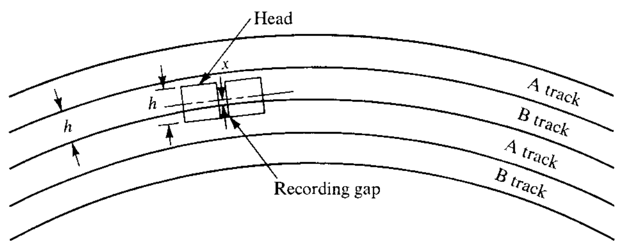

Figure 2 Servo disk format.

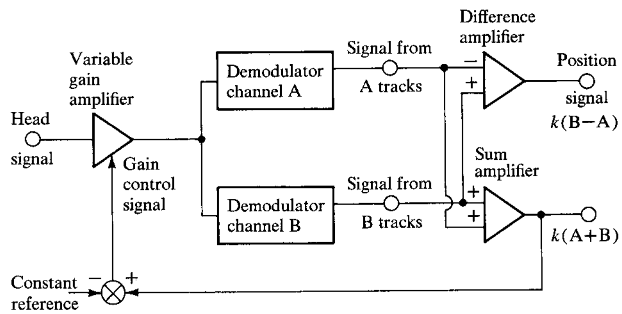

Figure 4 Electronic circuits for the demodulator-demultiplexer for the position transducer system.

unit of input radial misposition distance, should be a constant, independent of distance off track [3], head-signal amplitude variation, head-disk spacing variation, etc. These requirements may be accomplished by utilizing the property of a magnetic recording head in which the amplitude of the electrical read signal is proportional to the fractional width of the head over a recorded track. This is illustrated in Figs. 2 and 3(a) where the head width h is the same as the track width (see also Ref. [4]).

The servo head signals are processed in a way that a resulting signal, which we shall call the "position signal," is generated, and may be defined as

$$\text{Position signal} = k(B - A), \qquad (1)$$

$B \doteq$ portion of head signal due to track "B",
$A \doteq$ portion of head signal due to track "A",
$k \doteq$ gain parameter adjusted to make

$$k(A + B) = \text{constant}. \qquad (2)$$

Circuit configurations suitable for generating the position signal described by Fig. 3(b) and Eqs. (1) and (2) are shown in the block diagram of Fig. 4.

It is evident from Figs. 3 and 4 that the position transducer system consists of a modulator-multiplexer (the encoded servo tracks and servo head), communications channel (the head signal), and demodulator-demultiplexer (circuits of Fig. 4). The best encoding method for the servo disk includes the possible alternatives of frequency division multiplexing and time division multiplexing. Frequency division multiplexing would encode "A" tracks with one continuously recorded frequency and "B" tracks with a different frequency. The demodulator channels would then consist of appropriate band-pass filters, followed by envelope detectors. Time division multiplexing for the application might be better described as pulse position encoding, in which a synchronization character is coincidently recorded in both

Table 1 Encoding systems

System	*Advantages*	*Disadvantages*
Frequency division	1. Time synchronization between "A" and "B" tracks not required. 2. No need to regenerate sync information for demodulator. 3. Impulse noise rejection can be very good. 4. Demodulator is very simple.	1. Frequency-dependent amplitude variations cause off-track error. 2. Filter amplitude and frequency stability difficult to achieve. 3. May require ac biased recording to avoid wide harmonic spectrum in readback signal. 4. Index encoding usually adds considerable complexity to electronics.
Pulse position	1. Can utilize same recording and readback components as data channel. 2. Frequency response variations do not cause off-track error. 3. Demodulator sync directly usuable as data write oscillator sync. 4. Index encoding easy to add to existing sync information.	1. Sync information sensitive to impulse noise. 2. Sync recovery and gated demodulator electronics may be complex. 3. Accurate time synchronization required between recorded "A" and "B" tracks. 4. More sensitive to small disk-coating defects.

"A" and "B" tracks, followed by two mutually nonoverlapping time periods containing pulses only in "A" tracks or "B" tracks, respectively. A comparison of these two systems is made in Table 1.

In the initial design stages of the IBM 3340 servo surface encoding, practical, implementable solutions were quickly found for the problems identified in Table 1 as "disadvantages" of pulse position encoding. Much of the work done in implementing this class of encoding for the previously developed IBM 3330 disk file proved to be directly usable. On the other hand, certain solutions to the frequency-response and filter stability problems [disadvantages 1) and 2), Table 1] of frequency-division systems proved to be unattainable at competitive cost with electronics technology available in 1970 [5]. Therefore, a pulse-position servo encoding was chosen for use in the 3340, and engineering efforts were concentrated upon finding a particular pulse code that would both improve performance and reduce cost and complexity, compared to the system used in the 3330 [6]. The code shown in Fig. 5 was, to the author's knowledge, first proposed by Mueller [7]. It has become known as the tri-bit pattern for somewhat obvious reasons.

In reference to Fig. 5, synchronization information is contained in the leading edge 50 percent level of the sync pulse, which starts a monostable timing circuit to produce logic gates at "A" pulse and "B" pulse times. These logic signals are used to control identical gated peak detectors i.e., the "A" and "B" channel demodulators, respectively. The difference between the demodulator outputs is the position signal described by Eq. (1) and Fig. 3(b), while their sum is fed back as a gain control signal to satisfy Eq. (2). The variation of the servo head readback signal with radial displacements from $x = 0$ of Fig. 2 is shown in Fig. 5. This system has proven to be very effective in practice and its advantages result from the following features.

The "A" and "B" channel demodulator circuits are identical and, furthermore, the code is so constructed that off-track errors due to the electronic circuits can be caused only by differences between the "A" and "B" channels. This feature can exploit the inherent matching of components on an integrated circuit, yet does not require precise parameter values, which are very difficult to achieve. It can be seen from Figs. 3 and 4 that, to have the same slope of the position signal at every null point (which is required to maintain negative feedback in the servo loop), an inversion of polarity is necessary on every other track. This inversion may be accomplished in any system by swapping the outputs of the "A" and "B" demodulators; but in the tri-bit system, where the demodulators are identical, only simple logic gating of the "A" and "B" gate signals is required, and not a selectable analog gain of -1, which is a potential source of error and complexity.

Sync recovery with tri-bits requires only a level comparator for the -50 percent of the head signal because the sync pulse amplitude never varies as a function of off-track location. Because the spacing between sync pulses in this implementation was chosen to be two data-channel byte time intervals, the pulses are well separated and index encoding is easily accomplished by adding an additional sync pulse between the "A" and "B" pulses only at index time. Thus, index detection and the data-channel phase-locked oscillator reference can be obtained very accurately with only simple electronic circuitry rather than with the additional complex phase-locked oscillator required by most pulse-position encoding systems. The price paid for this feature is, of course, the allocation of a large percentage of channel capacity to sync information, but the servo transducer bandwidth required is so low that this is a very profitable tradeoff.

The only significant disadvantage found with the tri-bit system is a sensitivity to group delay distortion in any portion of the servo signal processing circuitry ahead of the demodulators. The most commonly encountered forms of group delay distortion cause asymmetrical pulse shapes in which the trailing edges of pulses are longer than leading edges. In the case of tri-bit servo encoding, this effect combines with any adjacent pulse interference to make the "A" pulse smaller and the "B" pulse larger, resulting in an offset error in the (B-A) position signal of Eq. (1). This error may be minimized by keeping the pulses well separated in time and by using signal processing components such as filters and amplifiers that are free from excessive group delay distortion.

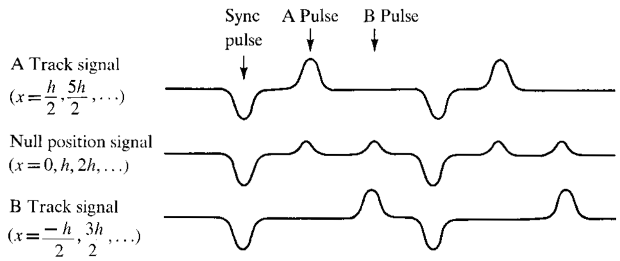

Figure 5 Readback signals for tri-bit encoding.

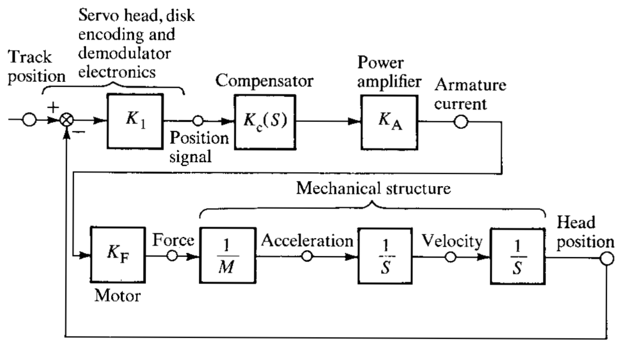

Figure 6 Block diagram for the track-following servo used in the IBM 3340. The units of the parameters are defined as: K_1 in volts/meter; $K_c(S)$ in volts/volt; K_A in amperes/volt; K_F in newtons/ampere; $1/M$ in meter/second^2/newton; $1/S$ in meter/second/meter/second^2 and $1/S$ in meter/meter/second.

Servomechanism dynamics

When the head-positioning servomechanism dynamics are analyzed, the structure, shown in Fig. 1, consisting of the heads and their supporting arms, carriage and bearings providing rectilinear motion, and the moving armature portion of the actuator motor, may be considered to be essentially an inertia, because both viscous and sliding friction have negligible effect. The type of actuator motor used, which has a moving coil armature attached to the carriage and a permanent magnet stator, is the so-called "voice coil motor," from analogy with permanent magnetic loudspeakers.

The force produced by such a motor is directly proportional to armature current, and the armature voltage has a component proportional to velocity, i.e., the "back emf," as well as the voltage drop due to current flow in the resistive and inductive impedance of the armature. Thus, the output impedance of the power amplifier used to drive the motor has a considerable effect upon system dynamics.

With a voltage-mode amplifier having a very low output impedance, the armature current, and thus the force, are both affected by the armature voltage, producing not only damping due to the back emf, but also a current rise time dependent upon the armature L/R time constant. On the other hand, a current-mode amplifier having a high output impedance is able to set the armature current independent of the armature voltage, within the limits imposed by its power supply voltages. Thus, with a current-mode driver, the current and force risetimes are much faster and are independent of the temperature-sensitive resistance of the armature, but there is no inherent viscous damping from back emf.

The advantage of a faster force risetime is very significant, and the loss of inherent motor damping can easily be made up by a compensation network in the servo electronic circuits. Nearly all designers choose current-mode drive, and the choice for the 3340 is no exception. A block diagram of the track-following servo of the 3340, using Laplace transform operational notation, is shown in Fig. 6.

The open-loop transfer function G of Fig. 6 is given by

$$G=\frac{K_c(S)\ K_1K_AK_F}{S^2M}. \qquad (3)$$

Without the compensator $K_c(S)$, (3) describes a classical second-order type-2 positioning servo [8]. It is capable of maintaining zero static misposition but, without

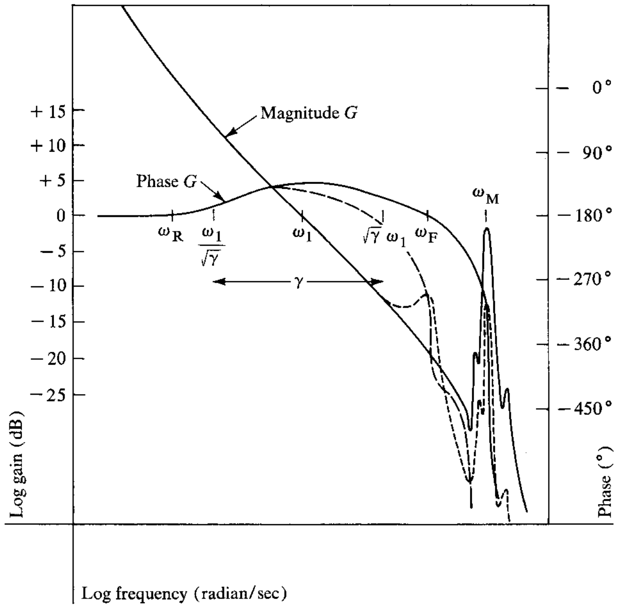

Figure 7 Bode diagram, showing how the upper limit on bandwidth of the closed-loop system is controlled by the need to maintain stability.

compensation, is a simple harmonic oscillator with no damping. The classical compensator for such a system is a lead-lag network $K_c(S)$ described by

$$K_c(S) = \frac{\omega_1^{\,2} M(1 + S\sqrt{\gamma}/\omega_1)}{\sqrt{\gamma} K_1 K_A K_F (1 + S/\sqrt{\gamma}\omega_1)} \qquad (4)$$

and substitution of Eq. (4) in Eq. (3) yields

$$G = \frac{\omega_1^{\,2}(1 + S\sqrt{\gamma}/\omega_1)}{\sqrt{\gamma} S^2 (1 + S/\sqrt{\gamma}\omega_1)}\,. \qquad (5)$$

The open-loop transfer function G of Eq. (5), which includes the compensator, is shown on the Bode diagram of Fig. 7 as an aid to seeing the effect of the compensator parameters γ and ω_1.

Parameter ω_1 is the radian frequency at which the open-loop gain is unity; it sets the bandwidth of the system. Parameter γ is the ratio of pole and zero frequencies $\sqrt{\gamma}\ \omega_1$ and $\omega_1/\sqrt{\gamma}$, respectively, of the compensator; it controls the size of the phase lead at ω_1 and thus the amount of damping in the closed-loop system. A value of $\gamma = 7.5$ yields a phase lead of about 50°. This value results in a linear system whose time response to a step displacement input has a single overshoot of less than 10 percent and is usually considered to be a most desirable settling characteristic.

The closed-loop system bandwidth is set by ω_1. Increasing bandwidth while holding other system parameters constant reduces the track-following error due to (1) radial runout of the tracks at the disk rotation frequency ω_R, (2) static forces produced by imperfections in the carriage bearings, or (3) cooling air circulating past the carriage. Increasing bandwidth also reduces settling time after a move between tracks and thus reduces access time, but the upper limit on bandwidth is controlled by the need to maintain stability in the presence of high-frequency mechanical-structural resonances, such as is shown in Fig. 7. If the bandwidth is increased sufficiently, the open-loop gain at frequency ω_M exceeds unity and most likely results in instability and oscillation. It is, of course, desirable to correct this by mechanical design improvements which lower the magnitude of the resonance or increase its frequency, but further improvements soon prove very costly.

In the 3340 track-following servo a sharp cutoff, second-order, active low-pass filter was added to the compensator to help attenuate this resonance and permit a larger, stable bandwidth. By placing the filter cutoff frequency ω_F at $5\omega_1$ and making its damping factor very small at $\xi = 0.25$, the peaking at ω_F was limited to about 6 dB and the phase loss at ω_1 to about −7°. By simultaneously increasing γ from 7.5 to 9, the small-signal system damping, transient overshoot, and gain margin at ω_M have remained nearly unchanged while increasing ω_1 (and thus the closed loop bandwidth) by 40 percent. There are at least two significant hazards in adopting this filter in other systems; it takes only a small-amplitude, unexpected structural resonance near ω_F to cause oscillation at that frequency and, secondly, the asymptotic slope of the magnitude function at high frequency is then −60 dB/decade. Only a small decrease in ω_M due to mechanical tolerances can thus greatly decrease the stability margin.

Before leaving the subject of compensators, a note should be added on hardware implementations. The final compensator and filter described above is shown in Fig. 8(a), together with a power amplifier having an input voltage offset V_g (an undesired error term) and transconductance of K_A amp/volt. The realization shown in Fig. 8(b) has the same small-signal transfer function as that shown in 8(a). However, with implementation (b) the offset voltage V_g, when referred to the position signal input, is clearly only $1/\gamma$ times as large as in (a), a reduction of nearly an order of magnitude for $\gamma = 9$!

Accessing control system

It has already been noted that the position transducer for the track-following servo, described by Figs. 3 and 4, is useful only over distances of about ±½ track and thus is incapable of controlling head motion over distances of many tracks. Furthermore, the objective of an accessing servo is to achieve minimum move time between any two tracks on the disk, rather than to achieve best accuracy in following a specific track. The control circuits of

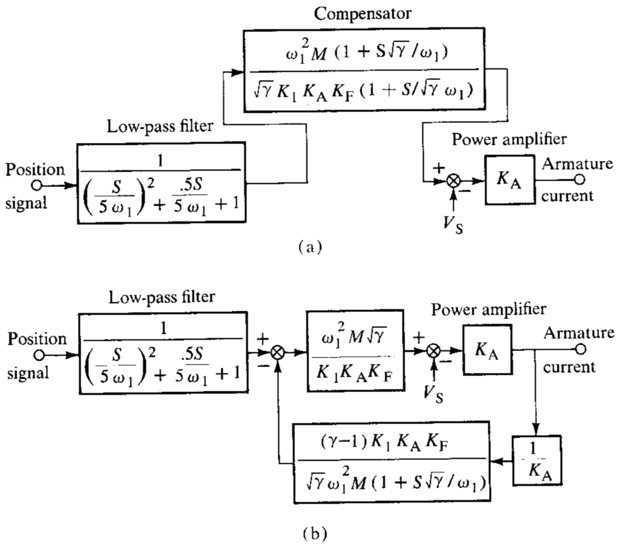

Figure 8 Track-following compensator. (a) As derived; (b) as implemented.

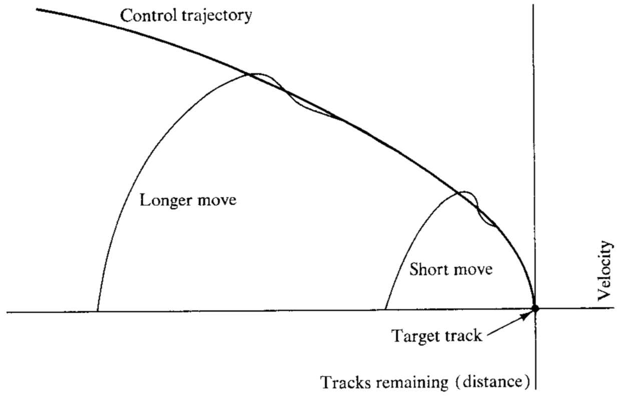

Figure 9 Access control trajectory.

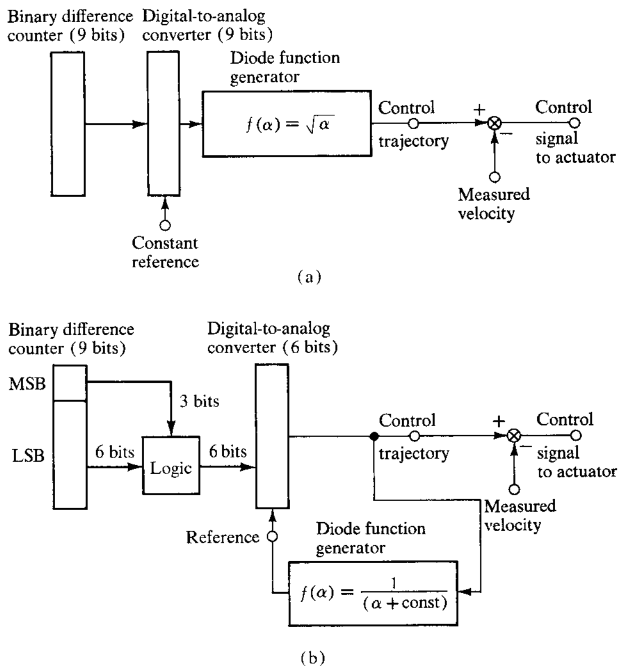

Figure 10 Method of generating the control trajectory. (a) Conventional method; (b) method utilized in the IBM 3340.

the accessing servo are thus much different from those of the track-following servo even though they both use the same power amplifier, actuator motor, and mechanical system components.

Time-optional control theory [1, 9] is utilized in the design of the accessing servo controller. In simple terms, the control strategy is to initially apply full forward power to the actuator until some point at which there is a switch to full reverse power until motion stops, hopefully at exactly the center of the desired destination track. However, implementation of such a system that also has the desired reliability is prohibitively expensive, and so a slight compromise is made in the control. Full forward power is still initially applied, but as soon as the system velocity and position correspond to a trajectory defined by the control circuits, reverse power is applied under closed-loop control. This maintains the system on the trajectory the remaining distance to the target track. Because the control trajectory is constructed in a way that a machine with worst-case performance can just follow with maximum reverse power, all other machines operate at less than maximum reverse power but still take no longer than the worst machine. They are then assured of reaching the correct target. For systems with negligible actuator motor inductance, the control trajectory is parabolic, of the form

$$\text{control velocity} = \text{constant} \times (\text{remaining distance})^{1/2}. \qquad (6)$$

Operation of this system is illustrated in Fig. 9.

Input commands to the accessing servo are in the form of a direction and an incremental track difference count. This difference is loaded into a down counter that is decremented by the zero crossings of the track-following position signal shown in Fig. 3. The track-position information is thus constantly available as a binary number in the counter, and velocity is measured by an analog electronic tachometer. A conventional way to implement the control trajectory is shown in Fig. 10, compared with the improved method used in the 3340.

The binary representation of a 350-track difference requires a 9-bit binary word, but the only portion of the control trajectory with any significant slope or curvature is the last 63 tracks. Using logic to force the digital-analog converter to full scale whenever any of the three more significant bits of the counter are ON will permit use of a much simpler 6-bit digital-analog converter without significant change in performance [10]. In addition, putting the diode function generator in feedback to the reference input of the digital-analog converter permits a curve fit to the required square root function with a series of curved line segments, rather than with the straight line segments obtained with the conventional cascaded function generator. This combination has not only permitted generation of a more accurate control trajectory (the error in the derivative of velocity with respect to position being much smaller), but also has reduced the number of segments required in the function generator and thus its cost and complexity.

Concluding remarks

The objective of this paper has been to present some of the engineering criteria for selection of a new final design for a disk file head-positioning servo. Most published literature in this field discusses how to implement an *a priori* design choice and then presents the resulting performance data. It is hoped that the methodology shown here of starting with the reasons that the design approach was chosen over its alternatives will stimulate the innovation of still better control systems in the future.

References and notes

1. C. J. Brown and J. T. Ma, "Time-Optimal Control of a Moving-Coil Linear Actuator," *IBM J. Res. Develop.*, **12**, 372 (1968).
2. M. Hiyane, Y. Inoue, and M. Kurusu, "Development of Linear Motion Actuator," *Fujitsu Scientific and Technical Journal* **23**, 59 (June 1972).
3. The transducer transfer functions should be a constant, independent of distance off track, at least within limits of $\pm\frac{1}{2}$ track pitch distance.
4. R. Tickell, "A High-Performance Position Control," presented to Joint Automatic Control Conference, University of Washington, Seattle, Washington, Aug. 17-19, 1966.
5. The outcome might well be different, however, with the linear integrated circuit technology available in 1974.
6. G. R. Santana, U.S. Patent 3,534,344, October, 1970.
7. F. E. Mueller, U.S. Patent 3,691, 543, September, 1972.
8. G. J. Thaler and R. G. Brown, *Analysis and Design of Feedback Control Systems*, McGraw-Hill Book Co., Inc., New York, 1960.
9. G. Leitmann, *An Introduction to Optimal Control*, McGraw-Hill Book Co., Inc., New York, 1966.
10. F. R. Hertrick, "Average Motion-Times of Positioners in Random Access Devices," *IBM J. Res. Develop.*, **9**, 124 (1965).

Received April 8, 1974

The author is located at the IBM General Products Division laboratory at San Jose, California 95193.